Computational model-based analysis of spatial navigation strategies under stress and uncertainty using place, distance and border cells

Yanran Qiu[1], Shiqi Wang[2], Jiachuan Wang[3], Wenyuan Zhu[4], Yuchen Cheng[5], Beste Aydemir[1] , Wulfram Gerstner[6], Carmen Sandi[6], **Gedi Luksys**[7]

[1]Ludwig-Maximilians University of Munich, Munich, Germany; [2]Scripps Research, San Diego, CA, USA; [3]National University of Singapore, Singapore; [4]ETHZ, Zurich, Switzerland; [5]Harvard Medical School, Cambridge, MA, US; [6]Brain Mind Inst., EPFL, Lausanne, Switzerland; [7]University of Edinburgh, Edinburgh, UK

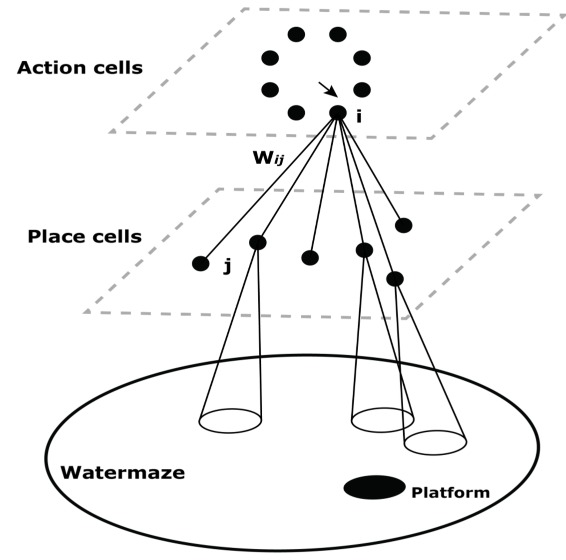
BACKGROUND

How do stress and uncertainty affect decision-making, learning and memory?

Decision-making occurs during navigation and learning. It is widely studied in choice behaviors, but less well understood in natural and more continuous settings, especially under **stress** and **uncertainty**. This process could be investigated in **rodent spatial navigation**, which has been modeled with place-cell-based models. However, traditional models usually ignored **detailed trajectories** or **kinematics**.

Here we extended a place cell-based reinforcement learning model to include detailed kinematics and used it to investigate the role of motivational stress in **Morris Water Maze**. We performed experiments with two **strains** of mice learning two versions of the task under different water **temperatures**: the task with a fixed platform location and the task where platform location varied randomly between two positions. Using **computational modeling** and **parameter estimation**, we were able to not only reproduce detailed mouse behaviors but also reveal **computational correlates of behavioral differences**. Our findings provide insights into computational mechanisms underlying spatial navigation in mice and how various modulators influence it.

PLACE-CELL-BASED MODEL



The position of the animal (state $s(t)$) is represented as a population activity of **'place cells'** (PC):

$$r_j^{pc} = \exp\left(-\frac{|\vec{p}-\vec{p}_j|^2}{2\sigma_{pc}^2}\right)$$

PCs project to a population of **'action cells'** (AC), representing directions of movement $\phi_i \in [0, 2\pi]$.

$$Q^{pc}(s_t, a_t) = r_i^{ac,pc} = \sum_j w_{ij}^{pc} r_j^{pc}$$

Weight update: Weights w_{ij} are initialized as uniformly distributed randoms from $[0, w_{mult}]$ and decay at a rate w_{noise} at each time step, and updated as to decrease the reward prediction error δ :

$$\delta = Reward(t) + \gamma Q_{t+1} - Q_t$$

Here, γ is temporal reward discounting factor.

Action selection: matching law with exploitation factor β (= "log softmax")

$$p(a|s) = \frac{Q^{pc}(s,a)^\beta}{\sum_{a_i \in A(s)} Q^{pc}(s,a_i)^\beta}$$

Eligibility trace keeps a (decaying) record of performed actions:

$e_{ij}^{pc} = \lambda e_{ij}^{pc} + r_i^{ac'} r_j^{pc}$, where λ is the eligibility trace decay rate.

Finally, weights are updated as follows: $\Delta w_{ij} = \alpha\, \delta\, e_{ij}$, where α is the learning rate.

To simulate realistic trajectories, acceleration constant (acc) is applied in the selected direction and velocities decay with time based on V_{decay}:

$$(v_x, v_y) = \left((v_x, v_y) + acc(\cos(\phi_{sel}) , \sin(\phi_{sel}))\right) V_{decay}$$

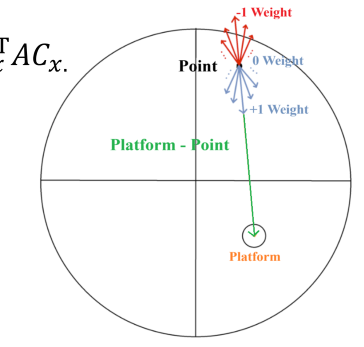
Once the modeled mouse hits the wall, only the tangential component is preserved.

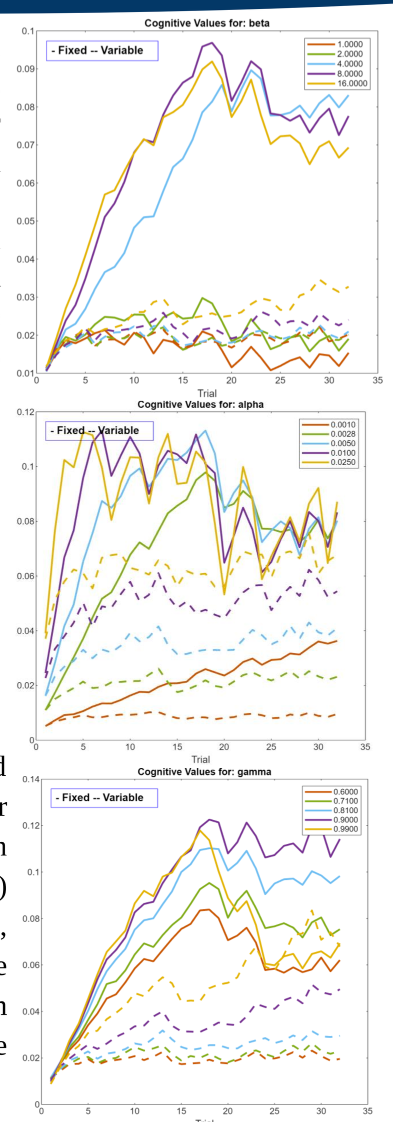
COGNITIVE VALUE

'Cognitive value' is a metric developed for assessing our model's ability to navigate in the water maze. The cognitive value of a location x represents the direction of the movement, which is a weighted sum calculated as the dot product of AC_x and the corresponding direction weights w_x:

$\rho_x = w_x^{\mathrm{T}} AC_x$.

Based on place cell model only.

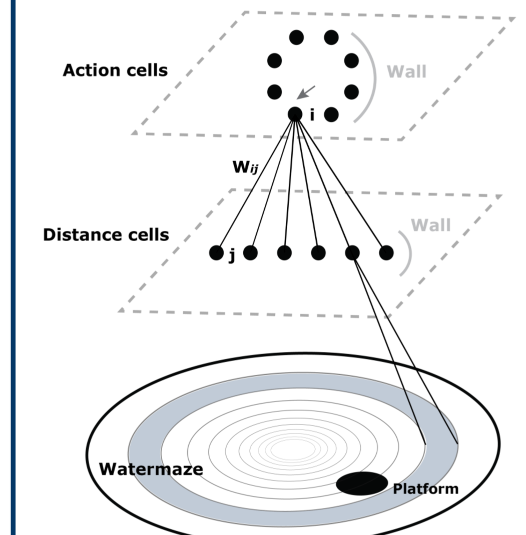




Higher beta (exploitation) produced more robust cognitive maps (higher cognitive values). Interestingly, high beta values (with beta value of 8 and 16) produced better immediate performance, while moderately high beta value (beta=4) produced better long-term performance. This is an example of the exploration-exploitation dilemma.

Cognitive values tend to be smaller for the variable platform.

MODEL EXTENSIONS



Wall distance-based cell model

'Distance cells' (DC) guide spatial learning via a cue-like signal, distance to the wall, which reproduced mouse behavior in tasks with uncertain platform positions better than place-cell-based strategies alone.

The PC-AC projections pass the absolute coordinates of mice in the water maze while DC-BC projections transmit the distance to the wall. BC-AC projections deliver both distance and angle.

While PC-AC alignments are fixed, DC-BC and BC-AC alignments are shifted circularly using wall as a reference.

$$r_j^{dc} = \exp\left(-\frac{(d-d_j)^2}{2\sigma_{dc}^2}\right)$$

$ratio_{pc}$ and $ratio_{dc}$ are manually set.

e.g., pure place cell model if $ratio_{dc} = 0$ and vice versa.

$$p(a|s) = \frac{Q^{pc}(s,a)^\beta}{\sum_{a_i \in A(s)} Q^{pc}(s,a_i)^\beta} \times ratio_{pc} + \frac{Q^{dc}(s,a)^\beta}{\sum_{a_i \in A(s)} Q^{dc}(s,a_i)^\beta} \times ratio_{dc}$$

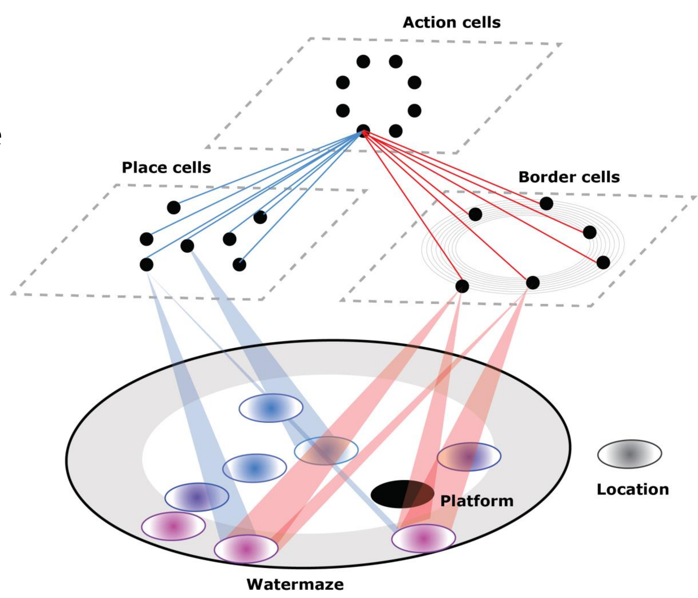
Place cell and border cell combined model

The border (boundary) cells are the cells with receptive fields distributed close to the border of water maze. They are sensitive to the distance to the wall like DCs when locations are within their receptive fields.

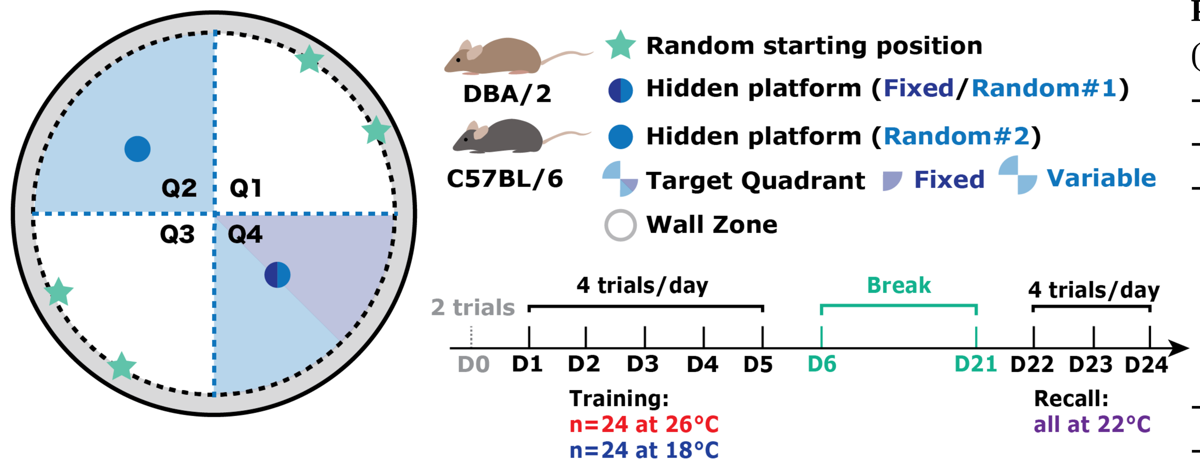
$$r_j^{bc} = \exp\left(-\frac{|d-d_j|^2}{2\sigma_{d_{bc}}^2}\right) \times \exp\left(-\frac{|\theta-\theta_j|^2}{2\sigma_{\theta_{bc}}^2}\right)$$

$$p(a|s) = \frac{Q^{pc}(s,a)^\beta}{\sum_{a_i \in A(s)} Q^{pc}(s,a_i)^\beta} \times ratio_{pc} + \frac{Q^{bc}(s,a)^\beta}{\sum_{a_i \in A(s)} Q^{bc}(s,a_i)^\beta} \times ratio_{bc}$$

$ratio_{bc} = 0.5 \times (1 - \frac{d}{r_{pool}})$ $\quad ratio_{pc} = 1 - ratio_{bc}$



EXPERIMENTAL SETUP

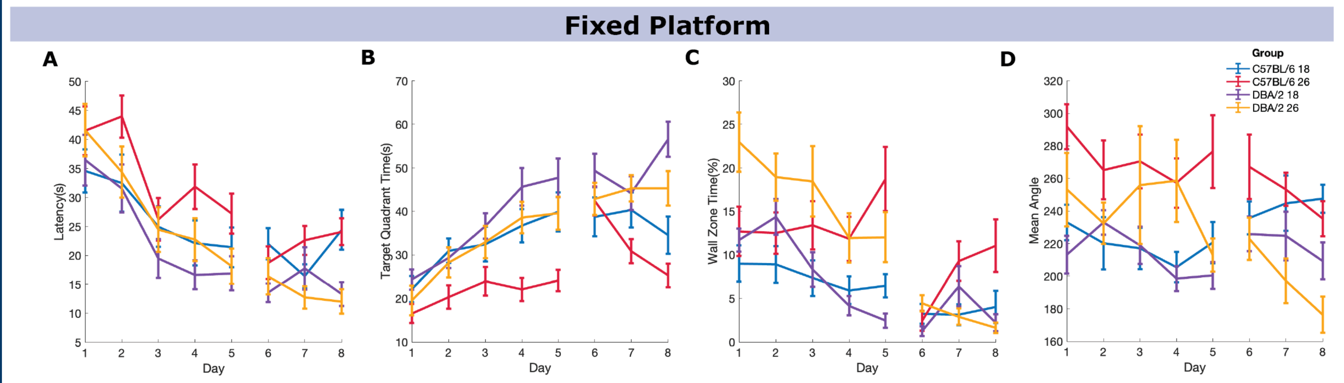


Performance measures (PM):

(for each mouse in each trial)

- Latency to platform (s)
- Swim distance (cm)
- Percentage of time (%) in
  - Target (possible) quadrant(s)
  - Opposite quadrant (fixed platform only)
  - Wall zone
- Mean turning angle
- Swim speed standard deviation (cm/s)

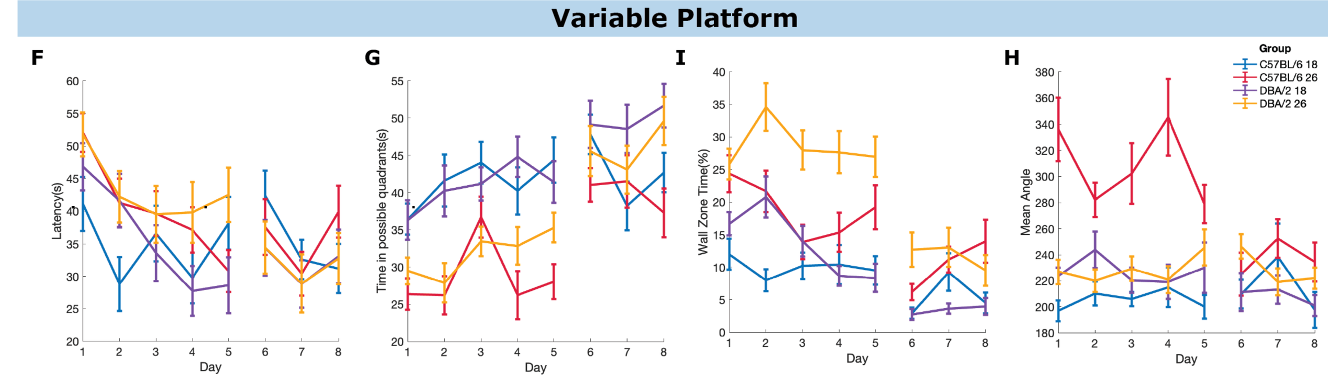
How do genetic predisposition and temperature stress influence performance on measures for learning and memory?



E

Training(Day1-5) Recall (Day6-8)

Mixed effect model: fix-strain, temperature, day, trial; random-mouse (*p≤0.05, **p≤0.01,***p≤0.001)								
p-value	latency		time in target quadrant		wall zone time		mean angle	
strain	**	***	**	***	0.096	**	*	***
temperature	0.085	n.s.	**	0.078	***	n.s.	***	n.s.
day	***	n.s.	***	n.s.	***	0.094	0.065	*
trial	*	n.s.	n.s.	*	n.s.	n.s.	n.s.	n.s.



J

Training(Day1-5) Recall (Day6-8)

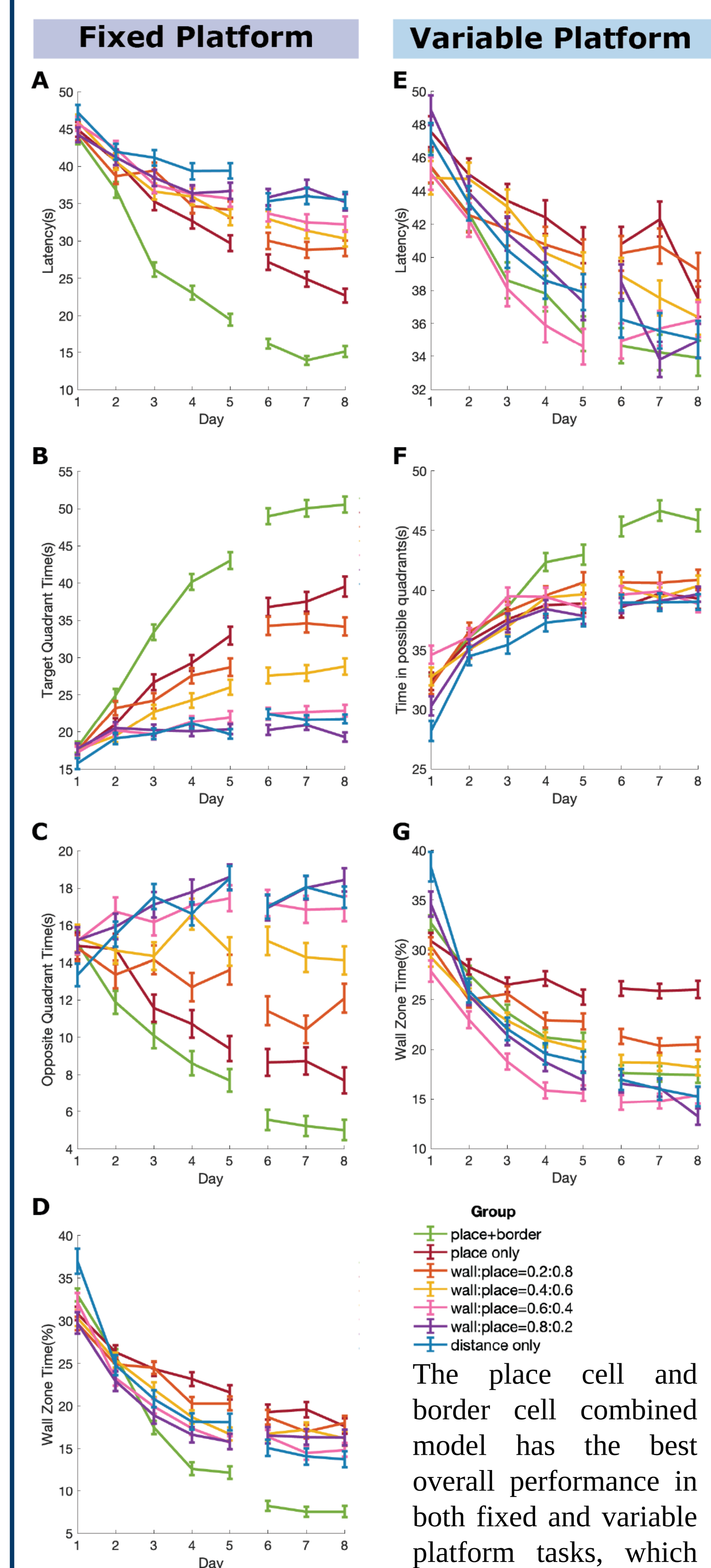
Mixed effect model: fix-strain, temperature, day, trial; random-mouse (*p≤0.05, **p≤0.01,***p≤0.001)								
p-value	latency		time in possible quadrants		wall zone time		mean angle	
strain	n.s.	n.s.	n.s.	***	***	n.s.	***	n.s.
temperature	**	n.s.	***	n.s.	**	*	***	*
day	***	n.s.	**	n.s.	***	n.s.	n.s.	n.s.
trial	n.s.	n.s.	0.068	n.s.	***	n.s	n.s.	n.s.

★ DBA/2 mice generally outperform C57BL/6 mice in the standard fixed platform task.

★ Mice at 26°C find the platform slower than those at 18°C for both tasks, but the difference disappears or becomes much less significant at the recall session. This indicates stress improves immediate performance, but not so much memory.

★ However, both mice trained at 26 °C perform much better when put to 22 °C water after the break (e.g., Time% in target quadrant for C57/BL 6 mice in the fixed platform greatly improved, better than they ever did before). This indicates that although they have been exploring at 26 °C they formed robust knowledge and used it under 22 °C.

MODEL COMPARISONS

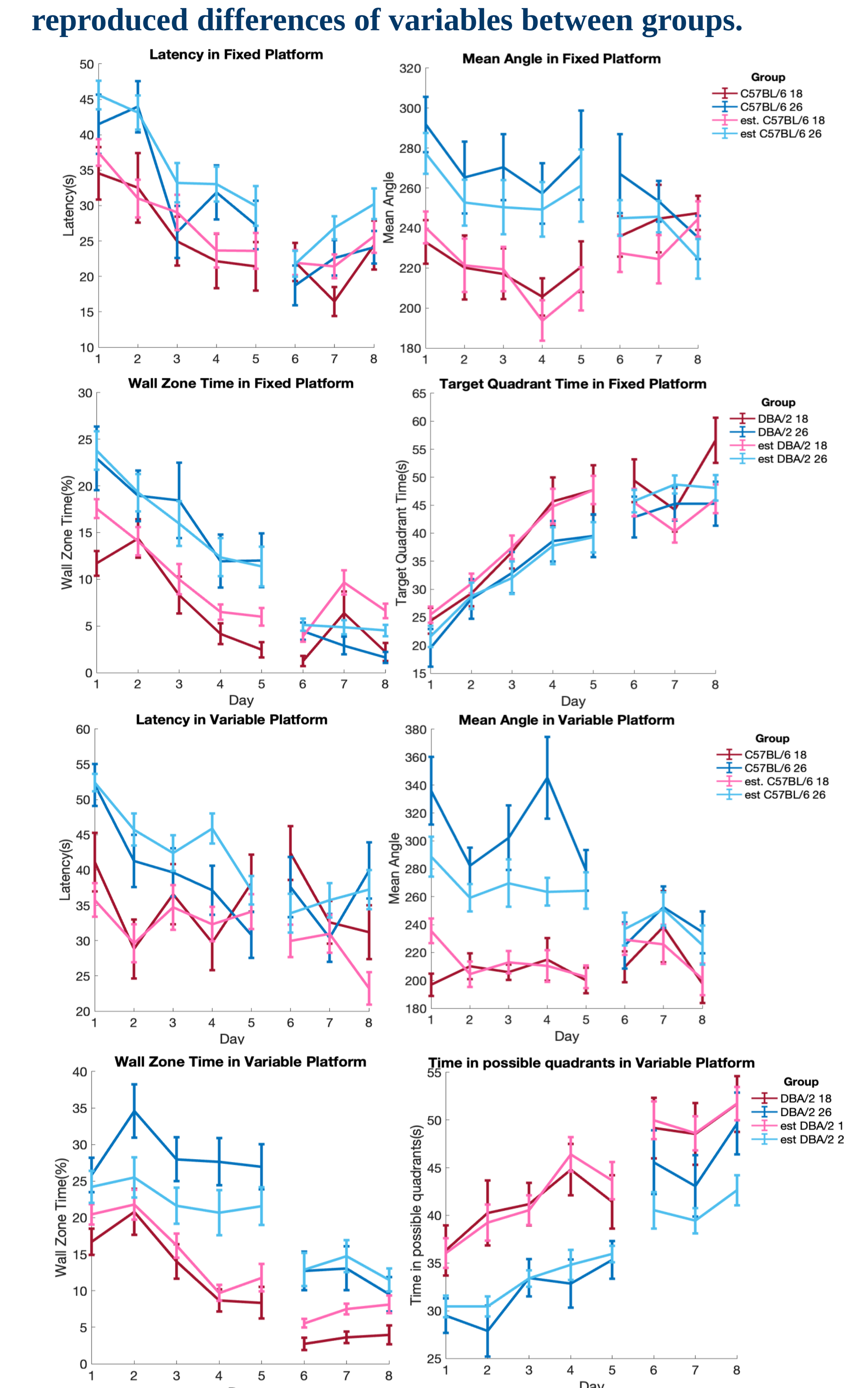


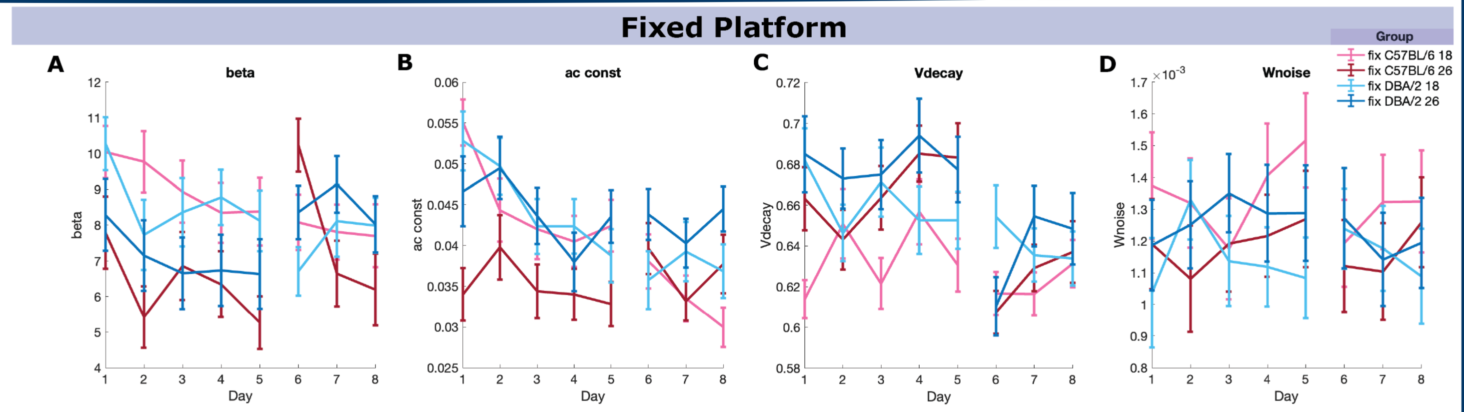
The place cell and border cell combined model has the best overall performance in both fixed and variable platform tasks, which has the best potential in fitting to animal behavioral data.

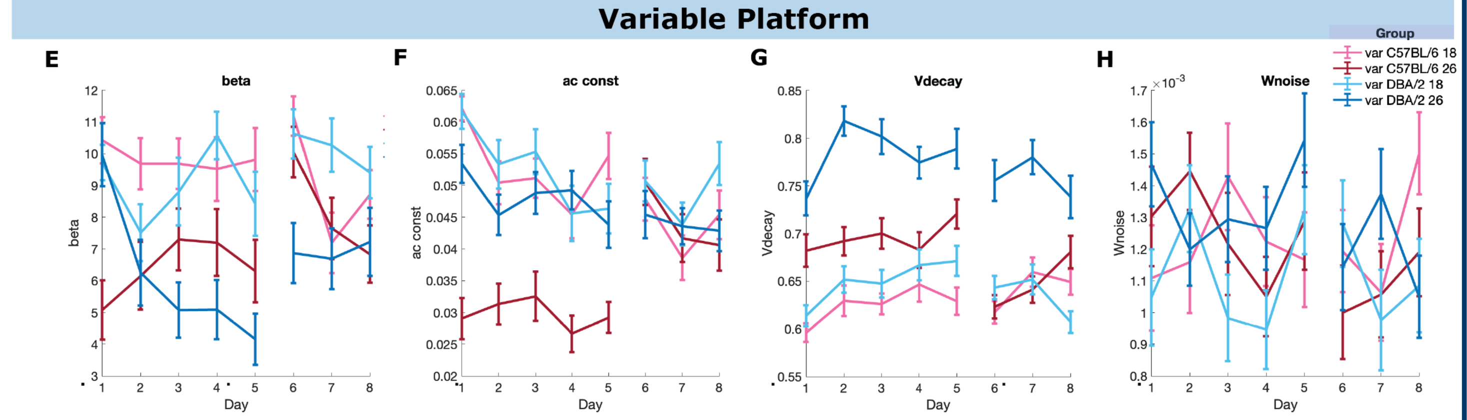
The comparison was done under same parameter setting.

PARAMETER ESTIMATION

We estimated best-fitting parameters in different experimental conditions and genetic strains of mice with the place and border cell combined model. We selected the least sensitive parameters to be fixed across days/conditions, as all cannot be flexible: **among them, crucially, the learning rate. The estimated fits reproduced differences of variables between groups.**







Model-derived interpretations

- Motivational stress induced by cold water improves spatial learning by modulating the exploration-exploitation balance (beta), rather than affecting the learning rate *per se*.
- Across tasks and strains, **cold stress consistently leads to increased exploitation of knowledge** (higher beta). These differences mostly disappear once the mice are put to 22°C water after the break.
- DBA/2 mice performing the variable platform task in warm water have considerably higher inertia of their movement (high V_{decay}), likely explaining their poor performance and high thigmotaxis (stay close to wall).
- C57BL/6 mice learning in warm water have considerably lower acceleration constants than other groups, especially in variable platform task, suggesting that whatever actions they choose (optimal or explorative), their performance vigour is reduced. This disappears once they are put to 22 °C water after the break.
- To further explore the relationship between model parameters and neuromodulators, experiments that manipulate neuromodulatory systems, especially dopaminergic and noradrenergic systems could be performed.

CONTACT

Email:
yanran.qiu@campus.lmu.de
gedi.luksys@ed.ac.uk

Twitter / X:
@yanran_qiu